

Point and Interval Estimation on the Degree and the Angle of Polarization. – A Bayesian approach –

D. Maier* and C. Tenzer, A. Santangelo

Institute for Astronomy and Astrophysics Tübingen (IAAT), Sand 1, 72076 Tübingen, Germany

Source: Publications of the Astronomical Society of the Pacific, Vol. 126, No. 939 (May 2014), pp. 459-468

Stable URL: <http://www.jstor.org/stable/10.1086/676820>

© 2014 by the Astronomical Society of the Pacific.

ABSTRACT

Linear polarization measurements provide access to two quantities, the degree (DOP) and the angle of polarization (AOP). The aim of this work is to give a complete and concise overview of how to analyze polarimetric measurements. We review interval estimations for the DOP with a frequentist and a Bayesian approach. Point estimations for the DOP and interval estimations for the AOP are further investigated with a Bayesian approach to match observational needs. Point and interval estimations are calculated numerically for frequentist and Bayesian statistics. Monte Carlo simulations are performed to clarify the meaning of the calculations.

Under observational conditions, the true DOP and AOP are unknown, so that classical statistical considerations – based on true values – are not directly usable. In contrast, Bayesian statistics handles unknown true values very well and produces point and interval estimations for DOP and AOP, directly. Using a Bayesian approach, we show how to choose DOP point estimations based on the measured signal-to-noise ratio. Interval estimations for the DOP show great differences in the limit of low signal-to-noise ratios between the classical and Bayesian approach. AOP interval estimations that are based on observational data are presented for the first time. All results are directly usable via plots and parametric fits.

Subject headings: polarization – confidence limits – Bayesian statistics – Methods: statistical – Methods: numerical

1. Introduction

Polarization of electromagnetic waves can be described with the concept of elliptical polarization, which implies linear and circular polarization as special cases. In this paper, the term ‘polarization’ is used to refer to linear polarization. Polarization measurements provide access to two quantities, the degree of polarization, which describes the relative amount of polarized photons to all observed photons, and the angle of polarization, which gives the orientation of the electric field of the electromagnetic wave. Point and interval estimations can be performed on both observables with the help of frequentist and Bayesian statistics.

Chandrasekhar (1950) rediscovered the Stokes parameters in 1950. Statistical considerations on the interpretation of polarimetric measurements are still an

frequentist				Bayesian			
degree of pol.		angle of pol.		degree of pol.		angle of pol.	
p.e.	i.e.	(p.e.)	i.e.	p.e.	i.e.	(p.e.)	i.e.

Fig. 1.— Point and interval estimations (p.e. and i.e.) for the degree and angle of polarization result in four key values that can be addressed by a frequentist and a Bayesian approach. The existing works by Simmons & Stewart (1985) (red), Naghizadeh-Khouei & Clarke (1993) (orange), and Vaillancourt (2006) (green) are labeled with colored boxes. The contribution of this work is colored in blue. Point estimations for the angle of polarization are trivial and are not mentioned explicitly in literature, see § 4.2.1 and § 5.2.1.

* e-mail: maier@astro.uni-tuebingen.de

ongoing point of discussion. Figure 1 summarizes the main achievements in the field to which this work refers. Serkowski (1958) emphasized that the observed degree of polarization is subject to biasing effects in a way that the observed value is preferentially greater than the true one. Simmons & Stewart (1985) presented several point estimators to correct for this effect. Furthermore, they constructed confidence intervals to determine the reliability of the estimated degrees of polarization. These works are based on frequentist statistics. Naghizadeh-Khouei & Clarke (1993) constructed confidence intervals for the angle of polarization as a function of the true value of polarization. These works rested on the assumption that the true value of polarization is known a priori; this is a fundamental concept in frequentist statistics.

However, in many situations with unknown true values, there is a need to construct point and interval estimations as functions of observational data. Based on a Bayesian approach, Vaillancourt (2006) proposed a method to construct credibility intervals that, in comparison to the classical confidence intervals, show significant differences in the region of low signal-to-noise ratios. Using Bayesian analysis, Quinn (2012) showed a complete summary of probability distributions for the degree and angle of polarization and possible priors suited to the experiment. The aim of this work is to extend the Bayesian approach to point estimations on the degree of polarization and to interval estimations for the angle of polarization.

§ 2 summarizes the foundations of the frequentist and the Bayesian approach. § 3 presents the underlying statistics that governs polarimetric measurements. Methods to construct point and interval estimations on the basis of frequentist statistics are reviewed in § 4 and recalculated with a Bayesian approach in § 5. This section includes a new method to choose the best estimator for the degree of polarization and a new method to construct interval estimations on the polarization angle. Monte Carlo simulated credibility intervals for the degree and angle of polarization are presented in § 6. Results and conclusions follow in § 7 and § 8.

2. Frequentist and Bayesian approach

The basis of frequentist statistics are probabilities of random events. Deductive reasoning leads to systems where unknown consequences can be studied for a known cause. In the example of polarimetric measurements, this means that the true degree and angle of

polarization of radiation that enters a polarimeter are known and the different possible manifestations within the observed data are investigated.

In contrast, Bayesian statistics extends the concept of probabilities to statements that become not only true and false, but more or less plausible. This reasoning allows one to deduce the plausibility of the cause on the basis of observed consequences. This is the usual case for all physical measurements: true values are estimated on the basis of observed data. In the case of polarimetric measurements, the true degree and angle of polarization are unknown parameters that shall be determined with the help of observed data.

Thus, frequentist and Bayesian statistics address two different points of view. A direct comparison of both methods is not reasonable because both methods result in different statements. To find the appropriate method for a given problem, one has to answer the question of whether the true values or the observables are known or unknown.

3. Statistics and Polarimetry

The normalized Stokes parameters q and u are appropriate variables to describe the linear polarization state of electromagnetic radiation (see Clarke et al. (1983) for more details). The degree and angle of polarization, P and Ψ , can be expressed as

$$P = \sqrt{q^2 + u^2} \quad \text{and} \quad (1)$$

$$\Psi = 0.5 \arctan(u/q). \quad (2)$$

Eq. (1) clearly shows that random noise σ in q and u results in a positively biased degree of polarization. $\sigma_q = \sigma_u$ is assumed. As the outcome of all type of polarimeters can be expressed in terms of Stokes parameters (Vaillancourt 2006), the following considerations are independent on the specific polarimeter type.

The applied terminology for a variable X is the following: x_0 labels the true value of X , \hat{x} is the estimated value, and x is the observed value. The main variables are:

P	degree of polarization
Ψ	sky-angle of polarization
p	signal-to-noise ratio of p (Eq.(3))
σ	error in q and u : $\sigma = \sigma_q = \sigma_u$
$\rho_{\Psi, (p)}$	probability density of Ψ (or of p)
$\rho(X A)$	probability density of a variable set X given the parameter set A

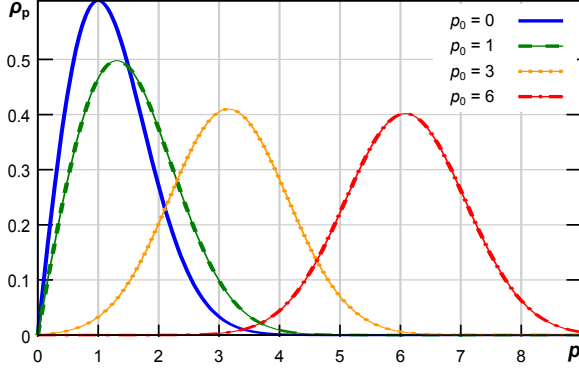


Fig. 2.— Probability density function ρ_p for different p_0 , see also Eq.(5). For large values of p_0 , the distributions become Gaussian shaped.

3.1. The signal-to-noise ratio of polarization

The probability densities used in the next sections depend on different parameters that determine their shape. These parameters are the degree of polarization of the source P , and the uncertainties of the measurement σ . They can be combined to derive the signal-to-noise ratio

$$p = \frac{P}{\sigma}. \quad (3)$$

The value of σ depends on the specific polarimeter type and observational conditions. See, for example, Elsner et al. (2012) for a calculation of σ for a counting based measurement with known background in the limit of low polarization.

With Eq. (3), the signal-to-noise ratio of any source observed with a specific polarimeter in a defined observation can be calculated. The signal-to-noise ratio p is independent from specific experimental conditions and serves as appropriate basis for the following statistical treatment.

3.2. The probability density function ρ

The detection principle of a polarimeter is to measure a sinusoidal variation of the signal¹ intensity as a function of the azimuthal angle. Based on that idea, Elsner et al. (2012) derived, in the case of low polarizations ($P^2 \ll 1$), the bivariate probability density function ρ which represents the probability of observing a signal-to-noise ratio p within the interval $[p, p + dp]$

¹The following calculations are based on the ideal case that observed polarimetric data are distributed continuously. Effects resulting from discontinuities, such as data binning, are not considered.

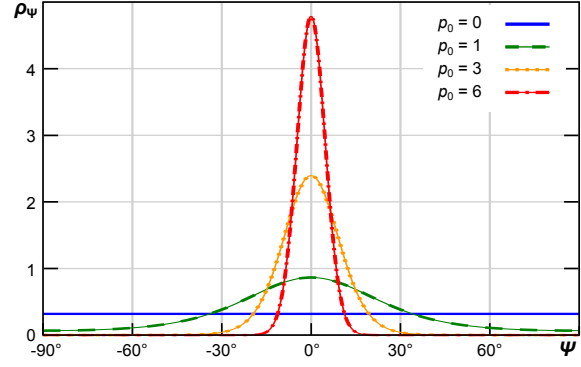


Fig. 3.— Probability density function ρ_ψ for different p_0 , see also Eq.(6). The distributions are non-Gaussian, but symmetrical around Ψ_0 (here: $\Psi_0 = 0^\circ$).

and a polarization angle Ψ within $[\Psi, \Psi + d\Psi]$ while the true values are p_0 and Ψ_0 :

$$\rho(p, \Psi | p_0, \Psi_0) dp d\Psi = \frac{p}{\pi} \exp\left(-\frac{(p - p_0)^2 + pp_0(2 \sin(\Psi - \Psi_0))^2}{2}\right) dp d\Psi. \quad (4)$$

Equivalent results² obtained under the assumption of Gaussian distributed Stokes parameters q and u , but without any limitations on p , are presented by Quinn (2012). Integrating Eq. (4), with respect to p , yields the Rice distribution (Rice 1945) – the univariate probability density function for the degree of polarization³ ρ_p (Elsner et al. 2012) – while integrating Eq. (4), with respect to Ψ , results in the univariate probability density function for the angle of polarization ρ_ψ (Naghizadeh-Khouei & Clarke 1993):

$$\rho_p(p | p_0) = p \cdot \exp\left(-\frac{p^2 + p_0^2}{2}\right) \cdot I_{1,0}(pp_0) \quad (5)$$

$$\rho_\psi(\Psi | p_0, \Psi_0) = \frac{1}{\sqrt{\pi}} \left(\frac{1}{\sqrt{\pi}} + \eta e^{\eta^2} (1 + \text{erf}(\eta)) \right) e^{-p_0^2/2} \quad (6)$$

where $I_{1,0}$ is the modified Bessel function of first kind and zeroth order, erf is the Gauss error function, and $\eta = p_0/\sqrt{2} \cdot \cos(2(\Psi - \Psi_0))$. The probability density functions are plotted in Fig. 2 and 3.

²This work is always related to the sky-angle and not the angle in the q - u -plane. Because $-\pi/2 < \Psi \leq \pi/2$ holds for Eq. (5) a factor of 1/2 is missing compared to other works that consider $-\pi < \Psi \leq \pi$.

³Statistical considerations on the degree of polarization (see § 4.1 and § 5.1) can be generalized to any fields in which a quantity is estimated from a quadrature sum of other quantities.

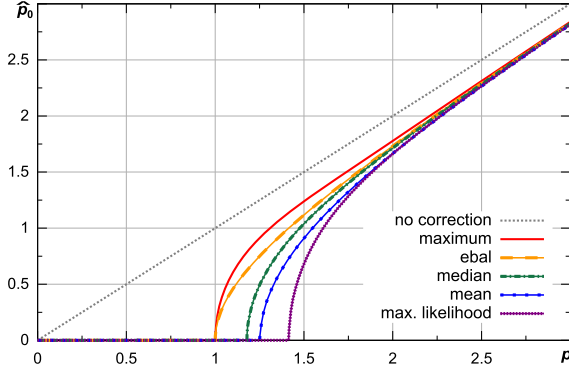


Fig. 4.— Plots of the maximum, median, mean, and maximum likelihood estimator presented in Simmons & Stewart (1985) in comparison with the ebal-estimator (Eq. (7)).

4. Point and interval estimation on the basis of frequentist statistics

4.1. Degree of polarization

The p -distribution ρ_p , in Fig. 2, shows that even unpolarized light ($p_0 = 0$) will yield an observed polarization $p > 0$. Simmons & Stewart (1985) showed several methods to correct for this biasing effect. They also presented a suggestion on how to construct confidence intervals around an estimated polarization. For completeness, we briefly review their results.

4.1.1. Point estimation

An estimator \hat{p}_0 can be constructed in different ways. Simmons & Stewart (1985) presented estimators based on the maximum, the median, and the mean of the $\rho_p(p|p_0)$ distribution. They also proposed an estimator based on the maximum of the corresponding likelihood function⁴ $\rho_{p_0}(p_0|p)$. In addition to this, we present a very simple estimator *ebal*⁵ that uses an approximate behavior of the maximum likelihood estimator for high signal-to-noise ratios (Vaillancourt 2006) in combination with a cut-off at $p = 1$:

$$\begin{aligned} \hat{p}_{0, \text{ebal}}(p) &= 0 & p \leq 1 \\ &= \sqrt{p^2 - 1} & p > 1. \end{aligned} \quad (7)$$

Polynomial fits for the other estimators are listed in Stewart (1991). Fig. 4 shows the graphs of all estimators.

⁴Cf. Eq. (13) and Eq. (14) and footnote 7.

⁵estimator based on approximated max. likelihood.

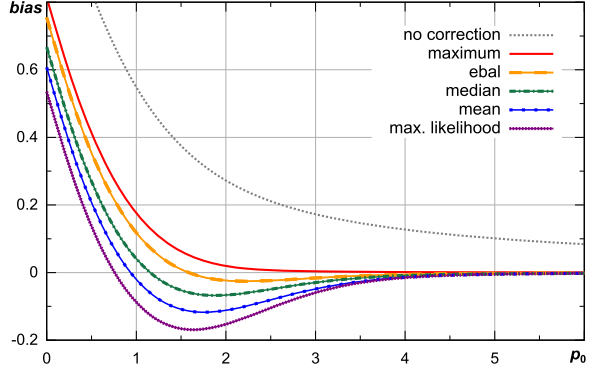


Fig. 5.— Bias of the estimators presented in Simmons & Stewart (1985) in comparison with the ebal-estimator (Eq. (7)). The max. likelihood estimator works best for low signal-to-noise ratios, whereas the maximum estimator works best for high values of p_0 .

tors. The estimated difference between the estimation value and the true value, the so-called *bias*, can be used to decide which estimator works best:

$$\text{bias}(p_0) = \int_0^\infty \rho_p(p|p_0) \cdot \hat{p}_0(p) dp - p_0. \quad (8)$$

The best estimator is the one with the smallest bias. Fig. 5 shows that all estimators work best in different regions of p_0 . Simmons & Stewart (1985) concluded that the maximum likelihood estimator should be used for low signal-to-noise ratios and the maximum estimator for high signal-to-noise ratios p_0 . This reasoning is not practically applicable because the estimators are used to obtain values for p_0 , so p_0 is unknown at the moment the best estimator must be chosen. See § 5.1.1 for a solution to this problem. Beside this problem, Simmons & Stewart (1985) did not consider the promising ebal-estimator, which has a bias smaller than 2.4 % for signal-to-noise ratios $p_0 > 1.4$, and also has a simple analytical description.

4.1.2. Confidence intervals

When the best estimator \hat{p}_0 for an observed polarization p is known, the confidence interval makes a statement about the reliability of this estimation. The definition of a confidence interval $\Delta p_0 = [\underline{p}_0, \overline{p}_0]$, with confidence level C , is related to a set of repeated measurements of an observable p for a true, but unknown, parameter p_0 . The probability that the constructed (\Rightarrow) confidence intervals, corresponding to the observables

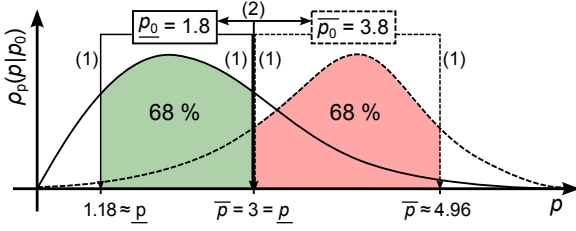


Fig. 6.— Illustration of the p_0 - p projection (Simmons & Stewart (1985)) for a σ_1 confidence interval in the case $p = 3$. If the true polarization is $p_0 = 1.8$ then 68 % of all observed polarizations p lie within $\Delta p = [1.18, 3]$. The corresponding interval for $p_0 = 3.8$ is $\Delta p = [3, 4.96]$. These calculations are based on Eq. (10). The σ_1 confidence interval for $p = 3$ is therefore $\Delta p_0 = [1.8, 3.8]$. See also Fig. 7.

p , contain the true value p_0 is C :

$$\forall (p \Rightarrow \Delta p_0) : \text{Prob}(p_0 \in \Delta p_0) = C, \text{ fixed } p_0. \quad (9)$$

In analogy with the Gaussian distribution, confidence intervals are labeled as σ_1 , σ_2 , and σ_3 for $C = 68.3\%$, $C = 95.5\%$, and $C = 99.7\%$. It should be noted that $\sigma_n \neq n \cdot \sigma_1$ as in the Gaussian case.

Confidence intervals make a statistical statement about the *intervals* containing the true value, but not about the *true value* being within a specific interval. The construction of a confidence interval is not a trivial calculation because it is related to the true polarization p_0 , but the distribution $\rho_p(p | p_0)$ (Eq. (5)) is a function of p . Simmons & Stewart (1985) constructed intervals $\Delta p = [\underline{p}, \bar{p}]$ by integrating ρ_p numerically, so that

$$\int_{\underline{p}}^{\bar{p}} \rho_p(p | p_0) dp = C \quad \left| \quad [\underline{p}, \bar{p}] \text{ minimal.} \quad (10)$$

The additional demand that Δp should be as narrow as possible is due to the fact that Eq. (10) is valid for a infinite number of intervals Δp . This demand, in combination with the asymmetric distribution ρ_p , results in intervals Δp that are not symmetric with respect to the maximum of ρ_p , as well.

Now, instead of an interval Δp obtained for a given p_0 , an interval Δp_0 for a given p is needed. This projection is done in two steps:

1. For all p_0 values corresponding \underline{p} and \bar{p} can be calculated with Eq. (10).

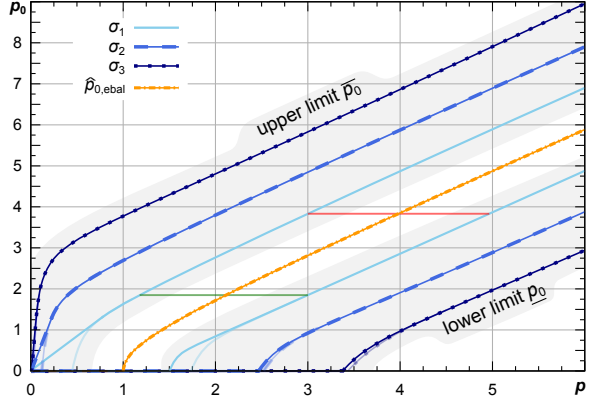


Fig. 7.— Confidence intervals $[p_0, \bar{p}_0]$ in terms of σ_1 , σ_2 , and σ_3 as function of the observed polarization p . The x/y-coordinates are interchanged compared to the similar plots shown by Simmons & Stewart (1985). The ebal estimator (orange graph in the center) shows that estimated values are in general not centered within the confidence interval. This becomes more prominent for small values of p and large confidence levels. The original construction method (Eq. (10)) is indicated in pale colors in contrast to the tangential progression (bold colors). The interval examples of Fig. 6 are plotted with consistent colors for illustrative reasons. Example: for an observed polarization of $p = 3$ the corresponding σ_1 -confidence interval is approximately $\Delta p_0 \approx [1.8, 3.8]$ while the estimated polarization is $\hat{p}_0 \approx 2.83$ by use of the ebal estimator. See Table 1 for parametric fitting results for \underline{p}_0 and \bar{p}_0 .

2. For a specific p value, only two p_0 - p - \bar{p} triplets are of interest: the one triplet with its upper limit \bar{p} equal to p defines the lower limit of p_0 (\underline{p}_0); and the other triplet with its lower limit \underline{p} equal to p defines the upper limit of p_0 (\bar{p}_0), see Fig. 6.

The following example may illustrate the situation for σ_1 confidence intervals (see Fig. 7 in pale colors): for $p = 3$, $\Delta p_0 \approx [1.8, 3.8]$ is well defined. For $p = 1.25$, the p_0 -projection (Fig. 6) is no longer applicable because the lower limit is not zero, but undefined. Nevertheless, the confidence limit $\Delta p_0 \approx [0, 1.9]$ is in accordance with Eq. (9), if it is assumed that Δp_0 does not exist for $p \lesssim 0.4$ (the intersection point of the upper limit curve with the p -axis).

To avoid the case of non-existing confidence limits, Simmons & Stewart (1985) departed from the term of narrowest Δp for small values of p_0 . They proposed a

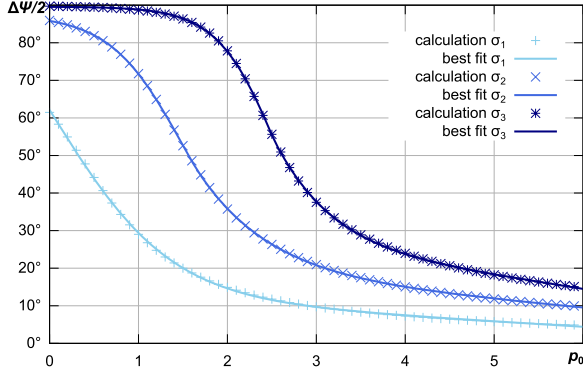


Fig. 8.— Uncertainties in Ψ in terms of σ_1 , σ_2 , and σ_3 confidence intervals as function of p_0 , (Eq. (11)). The plots were presented first by Naghizadeh-Khouei & Clarke (1993) (be aware of the wrong labeling in their Fig. 2a). The confidence intervals are $\Delta\Psi = [\Psi_0 - \Delta\Psi/2, \Psi_0 + \Delta\Psi/2]$.

tangential progression of the upper limit $\overline{p_0}$ to zero for $p \rightarrow 0$ and calculated corresponding new values of $\overline{p_0}$, so that the definition of Eq. (9) still holds. This way, confidence intervals can be set for all $p > 0$. Fig. 7 shows σ_1 , σ_2 , and σ_3 confidence intervals with and without the tangential progression. See Tab.1 for parametric fitting results.

Another example shows the interpretation of confidence intervals: for $p = 0.25$ the σ_1 confidence interval (by tangential construction) is $\Delta p_0 \approx [0, 0.4]$. This does not mean that the observed radiation is measured very precisely, but that a set of repeated measurements will result in different – most likely higher – values of p and new Δp that contain the true value p_0 in about 68 % of all cases.

4.2. Angle of polarization

4.2.1. Point estimation

Since there is no biasing effect in the angle of polarization (see Fig. 3), the trivial estimator $\hat{\Psi}_0(\Psi) = \Psi$ is sufficient.

4.2.2. Confidence intervals

Naghizadeh-Khouei & Clarke (1993) presented plots for $\Delta\Psi$ equal to σ_1 , σ_2 , and σ_3 confidence intervals $[\Psi_0 - \Delta\Psi/2, \Psi_0 + \Delta\Psi/2]$ by numerically integrating ρ_Ψ (see Fig. 8):

$$\int_{\Psi_0 - \Delta\Psi/2}^{\Psi_0 + \Delta\Psi/2} \rho_\Psi(\Psi | p_0, \Psi_0) d\Psi = C. \quad (11)$$

The basic idea in Eq. (11) is the same as in Eq. (10), with the only difference being that the symmetry of ρ_Ψ around Ψ_0 makes the narrowest confidence intervals symmetric around Ψ_0 , as well. The dependence of ρ_Ψ on p_0 is difficult because p_0 is unknown from an observational point of view.

Even though the plots of Fig. 8 are a function of p_0 and therefore not directly usable for observational data, it should be noted that the uncertainty of the polarization angle is a strictly monotonically decreasing function of p_0 . As Eq. (11) is independent of p , we see no way to construct confidence intervals for Ψ_0 as a function of p . See hereto the discussion in § 5.2.2.

5. Point and interval estimation on the basis of Bayesian statistics

So far, the frequentist approach – taking the true polarization p_0 as a fixed parameter – leads to the strange situation that the unknown true polarization p_0 must be known, or at least estimated, to select the best estimator \hat{p}_0 . Furthermore, the frequentist approach with its $p - p_0$ projection, results in a difficult construction of confidence intervals Δp_0 at low signal-to-noise ratios p and it seems impossible to construct usable confidence intervals $\Delta\Psi$ as a function of p .

All of these problems disappear with a Bayesian approach because now, the true polarization p_0 can be treated as a stochastic variable. In general, the posterior density $\rho_{p_0}(p_0 | p)$ can be computed with the likelihood $\rho_p(p | p_0)$ and the prior density $\rho(p_0)$ with Bayes theorem:

$$\rho_{p_0}(p_0 | p) = \frac{\rho(p_0) \cdot \rho_p(p | p_0)}{\int_0^\infty \rho(p_0) \cdot \rho_p(p | p_0) dp_0}. \quad (12)$$

In the following, we are considering a non-informative polar prior density: $\rho(p_0) = \text{const.}$ ⁶ Quinn (2012) studied the impact of non-informative prior densities, in general, and the difference of Jeffrey’s prior, which is uniform in the Stokes parameters q_0 and u_0 to the

⁶Technically speaking, this statement is critical because there is no uniform distribution living on the non-negative half-line \mathbb{R}_+ that can be normalized. This mathematical problem can be overcome by considering a maximal possible true degree of polarization $p_{0,\text{max}}$, so that $\rho = p_{0,\text{max}}^{-1}$. Finally: $p_{0,\text{max}} \rightarrow \infty$.

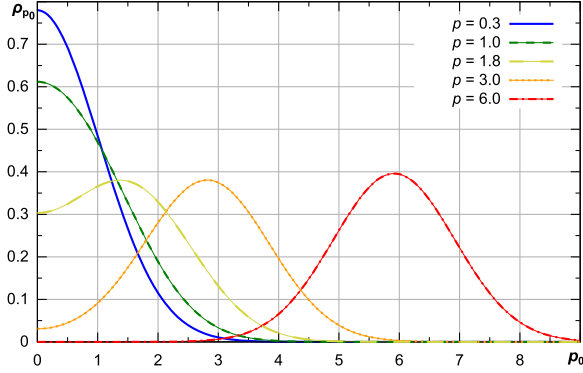


Fig. 9.— The posterior density distribution $\rho_{p_0}(p_0 | p)$ represents the probability density that an observed polarization p results from a true polarization p_0 . These distributions (Eq. (13)) were first calculated by Vaillancourt (2006).

uniform polar prior, which is uniform in p_0 . With reference to this work, it shall only be mentioned that Jeffrey's prior overstates large values of p_0 , so that the polar prior seems to be the best choice if any information from the source is missing:

$$\rho_{p_0}(p_0 | p) = \frac{1}{N} \cdot \rho_p(p | p_0) \quad (13)$$

$$\text{with } N = \int_0^\infty \rho_p(p | p_0) dp_0. \quad (14)$$

This is equivalent to a simple interchanging of parameter p_0 with variable p in Eq.(5), combined with a subsequent normalization⁷ N (Vaillancourt 2006). Fig. 9 shows a sample of posterior density distributions of ρ_{p_0} .

5.1. Degree of polarization

5.1.1. Point estimation

The probability density ρ_{p_0} can be used to transform any quantity that is a function of p_0 to an estimation value as a function of p . Using the already calculated $bias(p_0)$ of Eq. (8) leads to

$$bias(p) = \int_0^\infty bias(p_0) \cdot \rho_{p_0}(p_0 | p) dp_0. \quad (15)$$

⁷The normalization is unimportant for the calculation of the maximum likelihood estimator so that Simmons & Stewart (1985) simply set $\rho_{p_0} = \rho_p$ for their calculations.

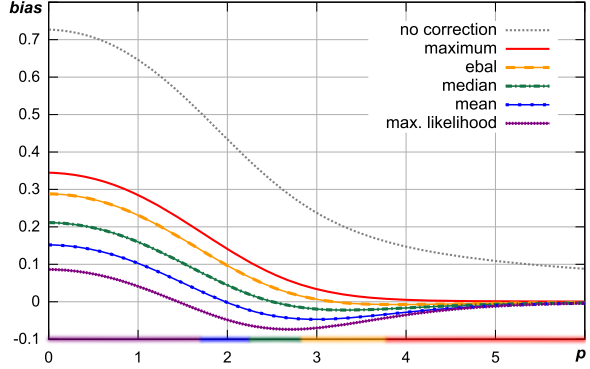


Fig. 10.— Bayesian *bias* as function of p . The maximum, median, mean, and maximum likelihood estimator are constructed like in Simmons & Stewart (1985) The ebal-estimator is defined in Eq. (7). The regions of best performance are indicated as color bars on the p -axis. See § 7.1 for listed values.

Each value of $bias(p_0)$ is weighted with the probability that this p_0 value results from an observed value p . Fig. The estimators in Fig. 10 show a similar behavior to those presented in Fig. 5, but now, the regions of the best performance are in terms of p . This allows the best estimator to be chosen directly for the first time on the basis of the observed data, cf. § 7.1.

5.1.2. Credibility intervals

Analogous to confidence intervals in frequentist statistics, credibility intervals in Bayesian statistics make a statement on the reliability of an estimated value. Different from confidence intervals, credibility intervals give the probability that the true value is within a specific interval. For all possible values of p_0 that can cause (\Rightarrow) the observed p , the credibility interval includes those p_0 that cause p in a fraction C of all cases:

$$\forall p_0 : \text{Prob}(p_0 \in [\underline{p_0}, \overline{p_0}] \Rightarrow p) = C, \text{ fixed } p. \quad (16)$$

Integrating the density distribution ρ_{p_0} over p_0 leads directly to credibility intervals (Vaillancourt 2006):

$$\int_{\underline{p_0}}^{\overline{p_0}} \rho_{p_0}(p_0 | p) dp_0 = C \quad \left| \quad [\underline{p_0}, \overline{p_0}] \text{ minimal.} \quad (17)$$

The resulting credibility limits are shown in Fig. 11 for σ_1 , σ_2 , and σ_3 credibility intervals. See Table 2 for parametric fitting results.

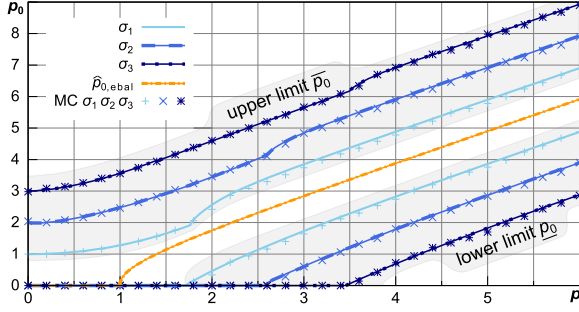


Fig. 11.— σ_1 , σ_2 , and σ_3 credibility intervals $[p_0, \bar{p}_0]$ as function of p . Example: for $p = 3$ the best performance is expected with the ebal estimator (see Fig. 10): $\hat{p}_{0,\text{ebal}}(3) \approx 2.83$. The corresponding σ_1 -interval is approximately $\Delta p_0 \approx [1.7, 3.8]$. The results of the Monte Carlo (MC) simulation of § 6 are plotted with points. See Table 2 for parametric fitting results for \underline{p}_0 and \bar{p}_0 .

5.2. Angle of polarization

5.2.1. Point estimation

As in § 4.2.1, the trivial estimator $\hat{\Psi}_0(\Psi) = \Psi$ is sufficient for point estimations on the angle of polarization.

5.2.2. Credibility intervals

Before proposing our method, we want to list the difficulties we recognized in constructing interval estimations $\Delta\Psi(p)$ and clarify why it is incorrect to proceed in the following way:

- Using the best estimator \hat{p}_0 as a parameter for Eq. (11) leads to incorrect results because the uncertainty of p_0 does not propagate into the uncertainty of Ψ .
- Calculating $\Delta\Psi$ with the lower limit \underline{p}_0 that was computed in § 4.1 overestimates the uncertainty in Ψ because the interval Δp_0 is constructed to be minimal in p_0 but not in Ψ (the upper limit \bar{p}_0 underestimates $\Delta\Psi$, cf. Fig. 8).
- Transforming p_0 to p , as in the transformation of the bias (Eq. (15)),

$$\Delta\Psi(p) = \int_0^\infty \Delta\Psi(p_0) \cdot \rho_{p_0}(p_0 | p) dp_0 \quad (18)$$

averages $\Delta\Psi$. The result is reasonable, but does not match the definition of credibility intervals in Eq. (16).

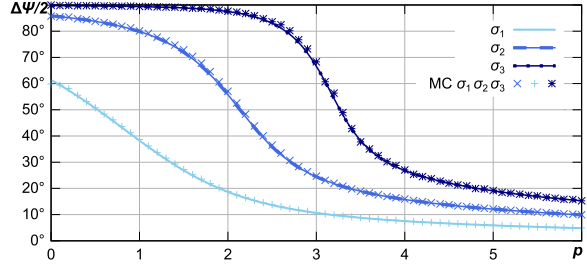


Fig. 12.— σ_1 , σ_2 , and σ_3 credibility intervals $[\Psi_0 - \Delta\Psi/2, \Psi_0 + \Delta\Psi/2]$. The results of the Monte Carlo (MC) simulation of Section 6 are plotted with points.

Our idea is to recalculate Δp_0 credibility intervals with the bivariate probability distribution $\rho(p, \Psi | p_0, \Psi_0)$ of Eq. (4). After integrating over p_0 and normalizing the bivariate distribution for fixed p 's with respect to Ψ , the credibility interval $\Delta\Psi$ can be computed (\Leftarrow) directly (without loss of generality $\Psi_0 = 0$):

$$\rho^*(p, \Psi) = \int_0^\infty \rho(p, \Psi | p_0, 0) dp_0, \quad (19)$$

$$\rho(p, \Psi) = \frac{\rho^*(p, \Psi)}{N(p)}, \quad N(p) = \int_{-\pi/2}^{\pi/2} \rho^*(p, \Psi) d\Psi, \quad (20)$$

$$\Delta\Psi(p) \Leftarrow \int_{-\Delta\Psi/2}^{\Delta\Psi/2} \rho(p, \Psi) d\Psi = C. \quad (21)$$

The numerically computed results are shown in Fig. 12. See Tab. 3 for parametric fitting results.

6. Simulating confidence intervals

The following Monte Carlo simulations serve as an illustration of the calculations made in § 5.

6.1. Degree of polarization

Our simulation is based on a large number N of (p_0/p) -pairs. For an equally spaced distribution of true signal-to-noise ratios p_0 , N_p p -values that follow the distribution of Eq. (5) are randomly chosen for each p_0 value. Our simulation uses $0 \leq p_0 \leq 10$ with a step size of $\Delta p_0 = 0.01$ and $N_p = 200\,000$. That makes $N = 2 \cdot 10^8$ (p_0/p) -pairs in total. To find (p_0/p) -pairs with a specific p -value, the continuous distributed pairs in p are binned in the p -dimension in intervals of $[p, p + 0.01]$, each containing N_{p_0} pairs.

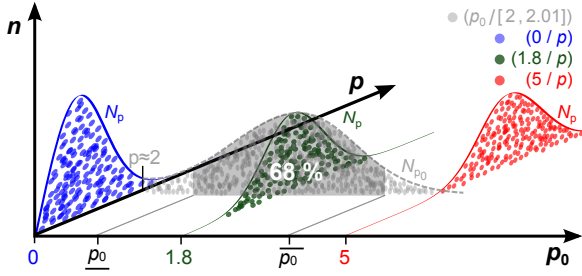


Fig. 13.— Model to explain the calculation of credibility intervals for the degree of polarization with the help of (p_0/p) -pairs: p -values are guessed on the basis of uniform distributed p_0 . The third dimension (n) equals the (p_0/p) -pair density. For reasons of clarity only the pairs with $p_0 \in \{0, 1.8, 5\}$ and those with $p = 2$ are shown. For the example $p = 2$, the corresponding σ_1 credibility interval (gray shaded area) consists of 68 % of all pairs $(p_0/2)$ with $p_0 \in [\underline{p_0}, \overline{p_0}]$.

The credibility interval $[\underline{p_0}, \overline{p_0}]$ for each p -value can finally be calculated as the narrowest interval in p_0 that contains $C \cdot N_{p_0}$ data points. Fig. 13 explains the described method graphically for $p \approx 2$.

6.2. Angle of polarization

Simulating interval estimations on the polarization angle ($\Psi_0 = 0$ is assumed) is done in two steps. First, appropriate random data must be simulated:

1. Choose a true signal-to-noise ratio $p_0 \in [0, 10]$.
2. Guess a value p following $\rho_p(p | p_0)$, cf. Eq. (5).
3. Guess a value Ψ following $\rho(p, \Psi | p_0, 0)$, Eq. (4).

Repeating these steps numerous times yields a large number of $(p_0/p/\Psi)$ -triplets. In the second step, analyzing these triplets will result in the desired credibility intervals:

1. Select all triplets with a specific value of $p \in [p, p + 0.01]$, independent of p_0 and Ψ .
2. Count the selected triplets $\rightarrow N_{p_0, \Psi}$.
3. Sort the triplet list with respect to Ψ .
4. Starting counting at $\Psi = 0$, the credibility interval⁸ $\Delta\Psi/2$ can be obtained as the Ψ value of the data triplet at list number $C/2 \cdot N_{p_0, \Psi}$.

⁸The Ψ -symmetry in Eq. (4) allows to restrict all calculations on the half credibility interval.

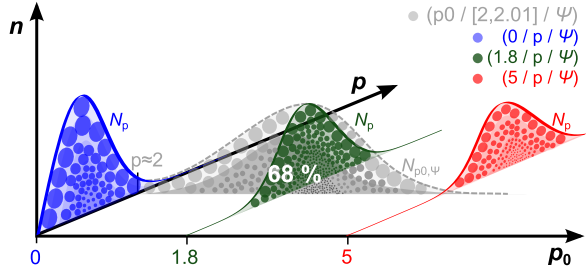


Fig. 14.— Constructing credibility intervals for the polarization angle. The notation is similar to Fig. 13 but now, each simulated data point is a triplet $(p_0/p/\Psi)$ plotted against its number density of occurrence n . As $\Psi_0 = 0$, it follows $\Delta\Psi = \Psi$. The values of Ψ are indicated by the size of the dots. For the example $p = 2$, the corresponding σ_1 credibility interval (gray shaded area) consists of 68 % of all triplets $(p_0/2/\Psi)$ with arbitrary p_0 but smallest Ψ .

In this way, the uncertainty in Ψ can be estimated on the basis of the observed signal to noise ratio p . Repeating these steps for a set of different p -values results in the data points plotted in Fig. 12. Again, the method used is explained graphically in Fig. 14.

7. Results

7.1. Degree of polarization

Based on the results shown in Fig. 10, the best estimator \hat{p}_0 can be chosen by means of lowest expected bias. The excellent results of the approximation $\hat{p}_{0, \text{ebal}} = (p^2 - 1)^{0.5}$ for $p > 2.8$ in combination with its analytical form makes this estimator a good choice for high signal-to-noise ratios p . The regions of the best performance and least square fits for \hat{p}_0 within those regions are the following:

$$\hat{p}_{0, \text{ml}} = 0 \quad p \in [0, \sqrt{2}] \quad (22)$$

$$\hat{p}_{0, \text{fit ml}} \approx (p - \sqrt{2})^{0.4542} + (p - \sqrt{2})^{0.4537} + (p - \sqrt{2})/4$$

$$\Delta = [-0.0078, 0.011], \quad p \in [\sqrt{2}, 1.70] \quad (23)$$

$$\hat{p}_{0, \text{fit mean}} \approx 22p^{0.11} - 22.076$$

$$\Delta = [-0.0039, 0.0035], \quad p \in [1.70, 2.23] \quad (24)$$

$$\hat{p}_{0, \text{fit med}} \approx 1.8p^{0.76} - 1.328$$

$$\Delta = [-0.0024, 0.002], \quad p \in [2.23, 2.83] \quad (25)$$

$$\hat{p}_{0, \text{ebal}} = (p^2 - 1)^{0.5} \quad p \in [2.83, \infty] \quad (26)$$

The Eq. (22)-(26) are not one combined estimator for all values of p , but a collection of different estimators,

each working in the region of best expected performance. Therefore, discontinuous jumps at the interval edges are not a lack of accuracy, but a result of different estimators. The motivation for the functional forms in Eq. (23) - (25) and in Eq. (27) and (28) are not physically driven but determined by mathematical intuition and the best fitting results. $\Delta = \hat{p}_{0,\text{fit}e} - \hat{p}_{0,e}$ indicates the range of maximal deviation between the fitted curves and the respective estimator e .

Concerning the construction of interval estimations, confidence and credibility intervals show very similar results for $p > 6$. At low signal-to-noise ratios $p < 3$ they differ significantly. The choice of which method to use depends on the question which shall be answered: is it the chance C that the confidence interval includes the true parameter, or the probability C that the true parameter is within the credibility interval?

Lower and upper limits for σ_1 , σ_2 , and σ_3 confidence intervals are presented in Fig. 7. A functional description for the tangential construction method is obtained by fitting the function

$$f(p) = Ap^B - Cp^{-D} + Ep \quad (27)$$

to the computed data points with the least square method. The results are listed in Table 1.

Table 1: Fitting results of Eq. (27) for the lower and upper confidence interval limits for the degree of polarization. See also Fig. 7. The deviation between the numerically calculated data points and f is maximal ± 0.05 . $f = 0$ for all undefined regions with $p < 6$.

f	A	B	C	D	E	validity
$\underline{p}_1(p)$	-1.17	0.027	5413	23.63	1.018	$1.51 \leq p \leq 6$
$\underline{p}_2(p)$	-1.86	0.192	$1.0e5$	13.58	1.083	$2.50 \leq p \leq 6$
$\underline{p}_3(p)$	-2.72	0.085	$1.0e7$	13.94	1.018	$3.41 \leq p \leq 6$
$\overline{p}_1(p)$	1.691	1.000	0.000	0.000	0.000	$0 \leq p \leq 0.77$
	0.715	0.328	0.021	4.766	0.936	$0.77 < p \leq 6$
$\overline{p}_2(p)$	5.441	1.000	0.000	0.000	0.000	$0 \leq p \leq 0.22$
	1.741	0.083	0.027	1.913	0.980	$0.22 < p \leq 6$
$\overline{p}_3(p)$	22.86	1.000	0.000	0.000	0.000	$0 \leq p \leq 0.065$
	4.621	-0.19	1.937	0.435	1.091	$0.065 < p \leq 6$

Calculated lower and upper limits for σ_1 , σ_2 , and σ_3 credibility intervals are presented in Fig. 11. Functional descriptions are obtained as before. The results

are listed in Table 2. For high signal-to-noise ratios $p > 6$, the Gaussian approximation can be used.

Table 2: Fitting results of Eq. (27) for the lower and upper credibility interval limits for the degree of polarization. See also Fig. 11. The deviation between the numerically calculated data points and the fit f is maximal ± 0.025 . $p < 6$ for all cases and $f = 0$ for all remaining undefined regions.

f	A	B	C	D	E	validity
$\underline{p}_1(p)$	4.241	1.021	2.286	1.134	-3.535	$p \geq 1.72$
$\underline{p}_2(p)$	0.468	1.177	3.974	0.874	0.145	$p \geq 2.54$
$\underline{p}_3(p)$	1.327	1.121	7.599	1.131	-1.00	$p \geq 3.45$
$\overline{p}_1(p)$	0.292	2.063	-1.00	0.000	0.000	$p \leq 1.72$
	0.855	0.020	17.87	6.012	1.009	$p > 1.72$
$\overline{p}_2(p)$	1.819	1.185	-2.00	0.000	-1.345	$p \leq 2.54$
	1.910	-0.028	13794	11.21	1.018	$p > 2.54$
$\overline{p}_3(p)$	0.564	1.632	-3.00	0.000	0.000	$p \leq 1.40$
	1.058	1.000	-2.47	0.000	0.00	$1.4 < p \leq 3.5$
	4.130	-0.38	$2.0e5$	10.65	1.140	$p > 3.50$

7.2. Angle of polarization

Despite its symmetrical probability distribution function ρ_Ψ , constructing confidence intervals for the angle of polarization is not trivial because of its dependency on p_0 . Calculated lower and upper limits for σ_1 , σ_2 , and σ_3 credibility intervals are presented in Fig. 12. A functional description is obtained by fitting the function

$$g(p) = A \left(B + \tanh(C(D - p)) \right) - Ep \quad [\text{in deg.}] \quad (28)$$

to the computed data points with the least square method.

For $p > 6$, a very simple description can be obtained using Gaussian error propagation on Eq. (1) and (2) with $\sigma_u = \sigma_q = \sigma$ (Serkowski 1962, Eq. (70)):

$$\sigma_1(p) = 28.65^\circ / p \quad p > 6 \quad (29)$$

$$\sigma_2(p) = 57.30^\circ / p \quad p > 6 \quad (30)$$

$$\sigma_3(p) = 85.95^\circ / p \quad p > 6. \quad (31)$$

Table 3: Fitting results of Eq. (28) for the lower and upper credibility interval limits for the angle of polarization. See also Fig. 12. The deviation between the numerically calculated data points and g is maximal $\pm 1.7^\circ$.

g	A	B	C	D	E	validity
$\sigma_1(p)$	32.50	1.350	0.739	0.801	1.154	$0.0 < p \leq 6.0$
$\sigma_2(p)$	65.65	0.323	0.858	2.688	0.000	$0.0 < p \leq 2.2$
	517	1.044	0.806	0.015	2.186	$2.2 < p \leq 6.0$
$\sigma_3(p)$	62.88	0.423	1.385	3.546	0.000	$0.0 < p \leq 3.2$
	102	1.380	1.327	2.506	3.958	$3.2 < p \leq 6.0$

8. Conclusions

Polarimetric measurements incorporate a non-Gaussian statistic in terms of the degree and the angle of polarization. The aim of this work was to present a systematic overview of the statistics that are necessary to analyze such measurements. In particular, we calculated point estimations for the degree of polarization and interval estimations for the angle of polarization with a Bayesian approach for the first time.

From an observational point of view, the Bayesian analysis shows substantial advantages compared to frequentist analysis. It allows direct access for the best estimator to be chosen on the basis of observational data and produces interval estimations with a meaningful interpretation.

In conclusion, observational polarimetric data can be recalculated in terms of signal-to-noise ratios p using Eq. (3). The choice of the best estimator, the best estimated value \hat{p}_0 , confidence and credibility intervals for the degree of polarization, and credibility intervals for the angle of polarization can then be obtained directly with the approximated formulas in § 7.1 and § 7.2. Using Eq. (3) reversed, all calculated values for the signal-to-noise ratio of the degree of polarization can be expressed as degree of polarization.

This work is supported by the *Bundesministerium für Wirtschaft und Technologie* through the *Deutsches Zentrum für Luft- und Raumfahrt e.V.* under the grant number 500O1110 and 50QR1101. We want to thank Prof. M. Zerner for discussing mathematical issues concerning confidence intervals.

REFERENCES

- Chandrasekhar, S. 1950, Radiative transfer. (Oxford, Clarendon Press)
- Clarke, D., Stewart, B. G., Schwarz, H. E., & Brooks, A. 1983, A&A, 126, 260
- Elsner, R. F., O'Dell, S. L., & Weisskopf, M. C. 2012, SPIE, 8443
- Naghizadeh-Khouei, J. & Clarke, D. 1993, A&A, 274, 968
- Quinn, J. L. 2012, A&A, 538, A65
- Rice, S. O. 1945, Bell Systems Tech. J., 24, 46
- Serkowski, K. 1958, Acta Astronomica, 8, 135
- Serkowski, K. 1962, in Advances in Astronomy and Astrophysics, Vol. 1, (Elsevier), 289 – 352
- Simmons, J. F. L. & Stewart, B. G. 1985, A&A, 142, 100
- Stewart, B. G. 1991, A&A, 246, 280
- Vaillancourt, J. E. 2006, PASP, 118, 1340